NeuroImage 124 (2016) 194–203

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Composition of complex numbers: Delineating the computational role of the left anterior temporal lobe

Esti Blanco-Elorrieta [a,*,c], Liina Pylkkänen [a,b,c]

[a] NYUAD Institute, New York University Abu Dhabi, Abu Dhabi, United Arab Emirates
[b] Department of Linguistics, New York University, New York, NY, USA
[c] Department of Psychology, New York University, New York, NY, USA

ARTICLE INFO

Article history:
Received 2 June 2015
Accepted 22 August 2015
Available online 30 August 2015

Keywords:
Conceptual combination
Magnetoencephalography
Left anterior temporal lobe
Numerical quantification
Complex number composition

ABSTRACT

What is the neurobiological basis of our ability to create complex messages with language? Results from multiple methodologies have converged on a set of brain regions as relevant for this general process, but the computational details of these areas remain to be characterized. The left anterior temporal lobe (LATL) has been a consistent node within this network, with results suggesting that although it rather systematically shows increased activation for semantically complex structured stimuli, this effect does not extend to number phrases such as 'three books.' In the present work we used magnetoencephalography to investigate whether numbers in general are an invalid input to the combinatory operations housed in the LATL or whether the lack of LATL engagement for stimuli such as 'three books' is due to the quantificational nature of such phrases. As a relevant test case, we employed complex number terms such as 'twenty-three', where one number term is not a quantifier of the other but rather, the two terms form a type of complex concept. In a number naming paradigm, participants viewed rows of numbers and depending on task instruction, named them as complex number terms ('twenty-three'), numerical quantifications ('two threes'), adjectival modifications ('blue threes') or non-combinatory lists (e.g., 'two, three'). While quantificational phrases failed to engage the LATL as compared to non-combinatory controls, both complex number terms and adjectival modifications elicited a reliable activity increase in the LATL. Our results show that while the LATL does not participate in the enumeration of tokens within a set, exemplified by the quantificational phrases, it does support conceptual combination, including the composition of complex number concepts.

Introduction

Understanding the brain basis of linguistic creativity is a fundamental goal for the cognitive neuroscience of language: what is the neurobiology of our ability to create an infinity of conceptual representations from the basic building blocks of language? Large networks of brain areas have been proposed to partake in the brain's "semantic network" (Binder et al., 2009; Binder and Desai, 2011) including the left inferior frontal cortex (e.g., Hagoort and Indefrey, 2014), the superior temporal gyrus (e.g., Friederici, 2011), the angular gyrus (e.g., Price et al., 2015) and the left anterior temporal lobe (LATL). Each of these regions has been proposed to carry a role in the combinatory processing of language. Damage to the angular gyrus can result in a wide variety of neuropsychological conditions affecting language, visuo-spatial processing and number cognition and thus it has been proposed as a high-level supramodal integration area, with the combination of concepts as part of its computational profile (Binder et al., 2009). The anatomical connectivity of the angular gyrus further conforms to a high level integrative role as it receives its input mostly from other association areas as opposed to primary sensory cortices (Bonner et al., 2013; Mesulam, 2000; Pandya and Seltzer, 1982; Yeterian and Pandya, 1985). The left inferior frontal cortex has also been associated with a multitude of functions, including phonological (Heim et al., 2008), semantic (Thompson-Schill et al., 1997) and syntactic processing (Stromswold et al., 1996), but within combinatory processing, its contribution has most commonly been proposed to be syntactic (Indefrey, 2012; Indefrey et al., 2001b; Hagoort and Indefrey, 2014; Friederici, 2011; Pallier et al., 2011; Tyler et al., 2011). Similar sensitivity to syntactic stimulus properties has been observed in posterior superior temporal cortex (Hagoort and Indefrey, 2014; Pallier et al., 2011).

However, as regards the semantic aspects of combinatory processing, multiple methodologies, including neuroimaging, electrophysiology and patient research, have produced an internally highly consistent body of work strongly implicating the LATL as a basic site for semantic combination. Core evidence for this include hemodynamic and neuropsychological research proposing that this brain area acts as a 'semantic

* Corresponding author at: New York University Abu Dhabi, Neuroscience of Language Lab, NYUAD Saadiyat Campus, A2, P.O. Box 129188, Abu Dhabi, United Arab Emirates.
E-mail address: eb134@nyu.edu (E. Blanco-Elorrieta).

http://dx.doi.org/10.1016/j.neuroimage.2015.08.049
1053-8119/

hub' in which conceptual representations are bound together and processed by a common set of neurons (Bright et al., 2004; Clarke et al., 2011, 2013; Gauthier et al., 1997; Grabowski et al., 2001; Rogers et al., 2006; Tyler et al., 2004) as well as sentence processing studies showing that structured sentences elicit greater LATL activity than meaningless sentences or word lists (Friederici et al., 2000; Humphries et al., 2006, 2007; Mazoyer et al., 1993; Pallier et al., 2011; Rogalsky and Hickok, 2009; Stowe et al., 1998; Vandenberghe et al., 2002; Xu et al., 2005). More recently, magnetoencephalography (MEG) studies on minimal combinations of two words have demonstrated that this activity relates to very basic combinatory operations as opposed to sentence-level phenomena both in comprehension (Bemis, and Pylkkänen, 2011, 2012) and in production (Del Prato and Pylkkänen, 2014; Pylkkänen et al., 2014).

While this large dataset on the LATL is still compatible with many definitions of "semantic processing", the robustness of these findings and their generality across multiple methodologies presents an opportunity for a systematic investigation of the computational details of this activity. One step towards sharpening our understanding involves recent MEG results on language production (Del Prato and Pylkkänen, 2014), where the modification of object denoting nouns with color adjectives (*blue cups*) engaged the LATL, while numerical quantification of the same nouns (*two cups*) did not. Given that both of these combinations involve semantic composition, these data are incompatible with a general semantic composition account of the LATL. Instead, they suggest a narrower computation, perhaps better characterized as a type of "conceptual combination", a label employed in the concepts and categories literature for a host of cases where, intuitively, the combination of two concepts serves to form a more complex one, typical examples being adjective–noun and noun–noun combinations. Given that in phrases such as *two cups*, *two* does not add a feature to the concept denoted by *cup* but rather enumerates the number of tokens in a set of cups, such cases would, by hypothesis, fall outside the definition of conceptual combination that is relevant for the LATL. Related evidence for the conceptual nature of the LATL include the sensitivity of its combinatory response to conceptual specificity (Westerlund and Pylkkänen, 2014; Zhang and Pylkkänen, 2015) and the correlation between the LATL activation elicited by specific concepts like *boy* and the product of the activations for their constituent concepts (i.e., *male* and *child*) (Baron and Osherson, 2011).

The purpose of the current experiment was to further characterize which input elements and specific computations constitute the "conceptual combinations" which drive activity within the LATL. Specifically, our study was designed around the question of whether complex number terms, such as *thirty-two*, would elicit combinatory activity in the LATL, despite its insensitivity to numerical quantification. Since this study builds on the results of Del Prato and Pylkkänen (2014), which was conducted in production, the current study is also a production study. Several prior studies have addressed the neurobiological similarity of combinatory operations in production vs. comprehension, with results compellingly showing that similar regions are recruited for composition whether the participant is comprehending or producing language (Hagoort and Indefrey, 2014; Menenti et al., 2011, Segaert et al., 2012; Pylkkänen et al., 2014). On the basis of this, one would predict the results of the current study to be replicable in comprehension. Further, as described in Methods, our production paradigm allowed us to keep the physical stimulus almost completely constant across conditions (cf., Del Prato and Pylkkänen, 2014; Pylkkänen et al., 2014), which was particularly useful given that confounding low level factors are often an issue in language studies. The combination of our two-word paradigm together with the millisecond time-resolution of MEG circumvents the principle obstacle behind electrophysiological investigations of sentence production, i.e., that meaningful electrophysiological data is extremely difficult to collect while the mouth is moving. However, the syntactic and semantic planning of small two-word phrases is thought to occur entirely prior to the onset of articulation (Alario et al., 2002; Meyer, 1996; Schriefers et al., 1999) and thus with a technique capable of capturing these planning stages millisecond-by-millisecond, we are able to measure combinatory processing (Pylkkänen et al., 2014). An added advantage of the detailed time resolution is that it allows us to separate different effects within the same region at different times.

Behavioral research on conceptual combination has classically been quite focused on one particular domain; the modification of nouns (e.g., Medin and Shoben, 1988; Murphy, 1990; Wisniewski, 1996; Hampton, 1997). Given that the LATL is at least activated by the core cases of conceptual combination, as evidenced by the many studies on adjective–noun combinations (comprehension: Bemis, and Pylkkänen, 2013; Westerlund and Pylkkänen, 2014 and production: Del Prato and Pylkkänen, 2014; Pylkkänen et al., 2014), it now becomes possible to concretely test what types of semantic combinations drive this activity. In other words, what is the brain's definition of "conceptual combination"?

Number words are a particularly interesting test case for this purpose as they are a very multifaceted word class in terms of the position and semantic functions they can fulfill in a sentence (Hurford, 1975). The most widely spread view states that (simplex) cardinals such as 'one', 'two' and 'three' are determiners (Barwise and Cooper, 1981; Bennett, 1975; Montague, 1974; Scha, 1984) and they have traditionally been treated either as generalized quantifiers (Montague, 1974; Barwise and Cooper, 1981) or restrictive modifiers (Link, 2002) when they precede the noun. However, according to Hurford (1975, 1987, 2001, 2003) and Ionin and Matushansky (2006): "when not acting as modifiers, the vast majority of simplex cardinals are singular nouns and belong to one or another open lexical class available in a language". Therefore, number words do not fall clearly in either open or close class word categories and can interestingly occupy the place of both in a noun phrase. This unique feature provides the opportunity to create different combinations and investigate to which extent the conceptual details of the input elements matter by creating a number of instinctively different combinations, while keeping the input elements constant.

The purpose of this experiment was to develop some understanding of the bounds and generality of the computations performed in the LATL regarding exactly what types of representations it combines. Particularly, as numerical quantification did not elicit conceptual combination in the LATL (Del Prato and Pylkkänen, 2014), our focus was on assessing whether this was because the LATL does not perform quantificational operations — which was Del Prato and Pylkkänen's interpretation — or because numbers in general are not a valid input to the LATL's combinatory mechanism. As a critical test case, we employed complex number terms such as *thirty-two*, which at least intuitively, may be instances of conceptual combination with numbers as the input. If such combinations engage the LATL while numerical quantifications do not, this would be evidence that it is the nature of the combinatory operation as opposed to the nature of the input items that matters for the LATL.

Like Del Prato and Pylkkänen (2014), our study employed a production paradigm where subjects named perceptually parallel displays in different ways, depending on task instruction. In all, our design included three combinatory conditions: complex number terms, numerical quantifications, and adjectival modifications, all of which were compared to non-combinatory list controls (Fig. 1). We aimed for minimal lexical differences in the produced utterances, and thus, given that complex number terms involve number words in both first and second position (*thirty two*), we designed the numerical quantifications to also have this property (e.g., *three twos*) while adjectival modifications involved a combination of a color adjective and a number term (*green twos*). As a primary non-combinatory control, we used lists consisting of two single-digit numbers (*two*, *three*), but also included lists consisting of a decade number and a single-digit number (*thirty*, *two*), given that lexically, this yields a form identical to the complex number term. However, given that decade numbers are themselves potentially complex, this latter control was not obviously non-combinatory, and thus could have been predicted to pattern somewhere between the combinatory conditions and our single-

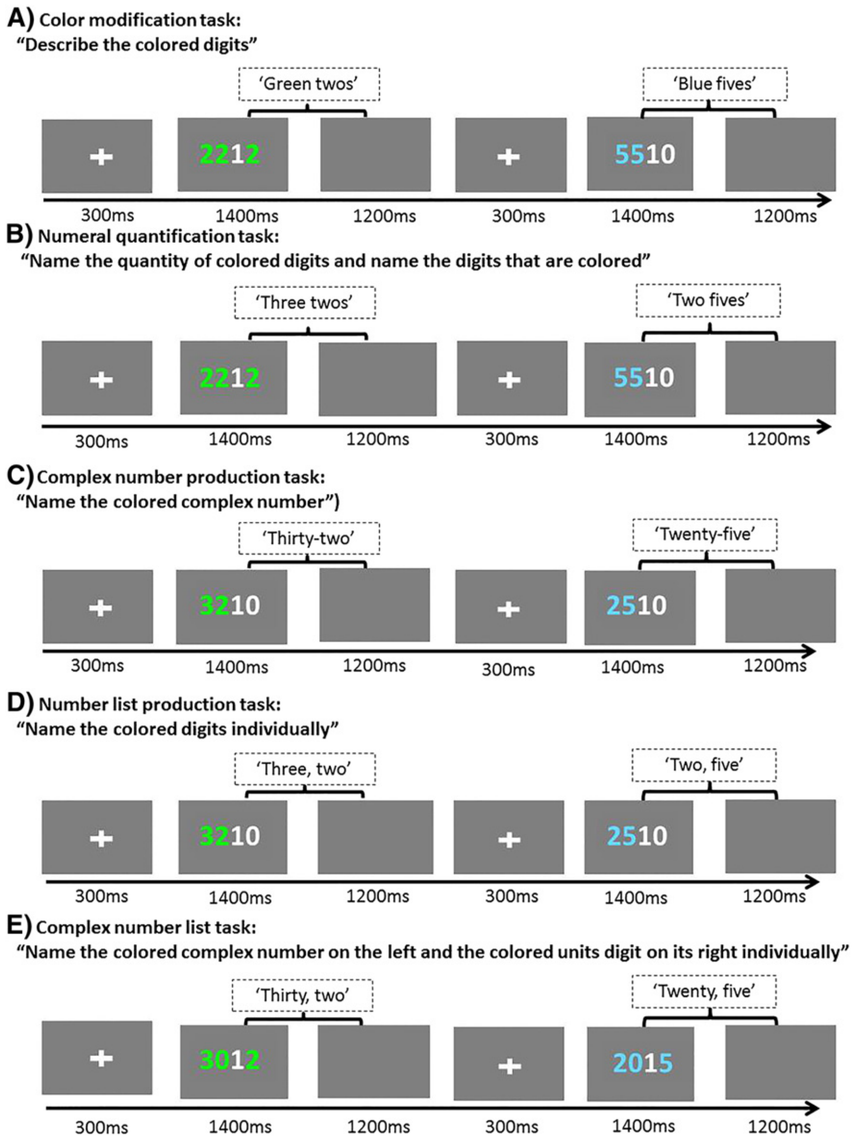

Fig. 1. Experimental design. Participants were asked to describe the colored digits onscreen by (A) naming the color and the numbers that were colored (B) counting the number of colored digits and naming the digits that were colored, (C) naming the complex number the colored digits formed, (D) naming the complex number on the left and the units number on the right individually or (E) naming the colored digits individually in a list like fashion.

digit list condition. In fact, this is what we observed (Additional Fig. 1). Thus, we decided to treat the straightforward number list (*two*, *three*) as the main non-combinatory control.

As depicted in Fig. 1, all our stimuli consisted of rows of four numbers, some of which were colored. In the Color Modification condition, participants were asked to name the color and the identity of the colored numbers as utterances such as "green twos". Although the cardinal name is not a prototypical noun, it is thought to act as a noun when placed as the head of the Noun Phrase (Hurford, 1975, 1987, 2001, 2003; Ionin and Matushansky, 2006). Our aim was to replicate the prior finding that adjectival modification of nouns engages the LATL (Bemis and Pylkkänen, 2011, 2013; Del Prato and Pylkkänen, 2014; Westerlund and Pylkkänen, 2014). Crucially, this condition assessed whether a non-canonical open class word such as a number word could ever elicit LATL-relevant conceptual combination.

In the Numerical Quantification condition, participants were asked to name aloud the number of colored digits and the name of those digits (e.g. "Three twos"). In this condition, the cardinal number which acts as a modifier characterized the exact cardinality of sets. Thus, in this case, 'three' meant "having exactly three members" (Bale and Khanjian, 2011; Ionin and Matushansky, 2006). In contrast, in the Complex Number condition, participants were asked to name aloud the two-digit complex number formed by the two colored digits onscreen (e.g., "thirty-two"). Crucially, both this condition and Numerical Quantification combined two number words, but the combinatory operation is different (quantifying the number of numbers vs. creating a complex number). On the basis of Del Prato and Pylkkänen's (2014) findings, we expected no LATL involvement in the processing of the numerical quantifications. If complex numbers patterned similarly, this would suggest that numbers in general cannot function as the "additional feature" whose incorporation constitutes LATL-relevant conceptual combination. In contrast, if complex numbers do engage the LATL, this would indicate that the combinatory operation housed in the LATL does not necessarily require conceptually rich input items, but instead, will also operate on featurally impoverished concepts such as numbers.

So far, we have assumed that at least potentially, the naming of complex numbers involves combinatory operations, but in fact, in the number literature, there has been significant debate about whether two-digit numbers are represented compositionally (i.e., each digit pair is processed as a decade digit and a unit digit separately; e.g. "thirty-two" is the combination of the concept "thirty" and the concept "two") (Grossberg and Repin, 2003; McCloskey, 1992; McCloskey et al., 1986; McCloskey et al., 1985; Nuerk et al., 2001; Verguts and De Moor, 2005) or holistically (i.e., each digit pair is processed as a single concept: "thirty-two" is stored as a single mental representation) (Brysbaert, 1995; Dehaene et al., 1990; Dotan et al., 2014; Reynvoet and Brysbaert, 1999). Evidence in favor of holistic representations has included studies reporting that priming effects are constant regardless of whether they involve priming by a single or two-digit number (e.g., the priming effect of 7 on 9 is the same as the priming effect of 11 on 9; Reynvoet and Brysbaert, 1999), studies finding that reaction times for deciding whether a two-digit number is larger or smaller than 65 show no significant discontinuities at decade boundaries (i.e., subjects respond "smaller" more slowly to 59 than to 51, although the 5 in the decades position already indicates that both of these numbers are smaller than 65; Dehaene et al., 1990), and studies showing that number magnitude, frequency of the number and sometimes the syllable length of the number name (but not differences in decade) influenced number reading times (Brysbaert, 1995). The main argument in favor of decomposition has, however, been that it is easier to judge whether a complex number is smaller than another when both the units and the decade digit are smaller (e.g. 67 and 52) than when only one of them is (e.g. 62 and 47) even if controlled for overall numerical distance, which suggests that subjects pay attention to the value of each digit, not the whole number (Nuerk et al., 2001; for more recent evidence see Verguts and De Moor, 2005; for a review see Grossberg and Repin, 2003). Additionally Zhou et al. (2008) suggest that whether complex numbers are processed holistically or compositionally depends on the stage of processing. In sum, then, whether complex numbers involve combinatory processing to begin with is an unsettled empirical question, and one that our dataset should be able to shed novel light on. It should be mentioned that even if the holistic vs decomposed dispute emphases on magnitude judgment, the "triple-code model" assumes that in addition to modality-specific symbolic codes in the visual Arabic and auditory verbal domain, there is also a supramodal abstract "number sense" that conveys semantic information (Dehaene, 1992). Thus, even if the task itself does not explicitly require processing of numerical magnitude this abstract "number sense" remains activated.

A final noteworthy aspect of our design was that both our Color Modification and Numerical quantification conditions involved the production of plural noun phrases (i.e., "blue threes" and "two threes"), whereas the complex number condition did not ("twenty-three"). Previous studies have found increased activation in the left angular gyrus when contrasting plural nouns that were morphosyntactically marked for the number feature (-s) with singular nouns (Domahs et al., 2012) and when linking back two-sentence discourses to plural rather than singular subjects (Boiteau et al., 2014). Given that the left angular gyrus (AG) has also been identified as the most likely candidate for number case agreement violations (Carreiras et al., 2010) and proposed to support the manipulation of numbers in verbal form together with other left perisylvian areas (Dehaene et al., 2003), we analyzed whether the angular gyrus would show effects of morphosyntactic plurality. For completeness, in addition to regions where LATL combinatory effects have localized in prior studies (Brodmann areas 38, 20 and 21), our analysis also included the ventromedial prefrontal cortex, a common though somewhat less consistent locus of basic combinatory effects (e.g., Bemis and Pylkkänen, 2011; Brennan and Pylkkänen, 2012) as well as the left inferior frontal gyrus (LIFG), given its traditional association with language production, though not, at least in MEG, basic combinatory processing either in comprehension (Bemis and Pylkkänen, 2011, 2012) or production (Pylkkänen et al., 2014).

Materials and methods

Participants

28 right-handed, native English speakers participated in this experiment (19 female 9 male, 24.7 years average 4.15 SD). All were neurologically intact, with normal or corrected-to-normal vision and all provided informed written consent.

Stimuli and experimental design

The experiment consisted of 480 trials in which participants were presented with four digits in a row. The numbers were presented on a dark gray background and some of the digits were colored in pink, blue or green while the rest were presented in light gray (e.g., participants would see "3218" where "3" and "2" were green and "1" and "8" were light gray). The stimuli were kept constant across all conditions, thus assuring that there was no perceptual variation amongst them. Participants were required to name the numbers colored in pink, blue or green aloud; the specifics of the naming task varying upon instruction. Participants were asked to either name the colored digits individually ('three, two'), the quantity of colored digits and the digits that were colored ('three twos'), the color of the digit and the digits that were colored ('green twos'), the complex number that they formed ('thirty-two') or the complex number and the units separately ('thirty, two') (Fig. 1). The quantity of colored digits and their location on the number varied in a controlled fashion between conditions and all numbers were composed by four digits to assure that participants could perceive all the digits of each number at one glance (Kaufman et al., 1949; Saltzman and Gamer, 1948).

In order to assure that participants were performing a genuine quantification task in the Numerical Quantification condition, we aimed for the base number word (which corresponded to the described colored number) to be bimorphemic (one morpheme in root, and -s). As a consequence, 'one' was excluded as the first-position word and only 'two', 'three' or 'four' were included. The color words for the color modification condition were chosen such as to match the quantifier number words as closely as possible while also being maximally visually distinctive from each other (if the latter constraint was not met, then the condition involving color naming could have turned out harder than number naming). Consequently, we required the color words to be monosyllabic like the number words. For the rest of lexical-level variables, we chose English Lexicon Project naming times as a summary statistic (Balota et al., 2007). The color words 'blue', 'pink' and 'green' were chosen as optimally matching these number words. The English Lexicon Project naming times of these color words were somewhat faster (mean = 577 ms) than those of the Number words (mean = 605 ms), but this difference was not significant ($p = .13$). In addition, only monosyllabic number words from 1 to 9 were included in the design as base numbers. We excluded disyllabic number words (e.g. "seven") to avoid effects related to larger number of syllables (such as delayed naming latencies and greater motor preparation for more syllabled words). Therefore, only eight numbers were used as base numbers ('one', 'two', 'three', 'four', 'five', 'six', 'eight', and 'nine').

All base number words were combined with the three possible first position words (either number words or colors) eliciting 24 combinations for each condition which formed each of the experimental blocks. Each block was repeated four times during the experiment, eliciting 96 trials per condition (480 trials in total). All pictures were presented foveally using Presentation (Neurobehavioral System Inc., California, USA) and subtended in a range from 55.16° height and 33.36° width on a screen ~85 cm from the subject.

Procedure

Before recording, each subject's head shape was digitized using a Polhemus dual source handheld FastSCAN laser scanner (Polhemus, VT, USA). MEG data were collected in the Neuroscience of Language Lab in NYU Abu Dhabi using a whole-head 208 channel axial gradiometer system (Kanazawa Institute of Technology, Kanazawa, Japan) as subjects lay in a dimly lit, magnetically shielded room. Vocal responses were captured with an MEG compatible microphone (Shure PG 81, Shure Europe GmbH).

In all conditions, trials began with a fixation cross (300 ms), followed by the presentation of the stimuli. The picture remained onscreen until speech onset (1400 ms timeout), and participants were allowed 1200 ms to finish their speech before the fixation cross for the following trial would appear. The entire recording lasted ~25 min.

Data acquisition and preprocessing

MEG data were recorded at 1000 Hz (200 Hz low-pass filter), epoched from 200 ms before to 700 ms after picture onset and noise reduced via the Continuously Adjusted Least-Squares Method (Adachi et al., 2001), in the MEG Laboratory software (Yokogawa Electric Corporation and Eagle Technology Corporation, Tokyo, Japan). All trials exceeding the absolute threshold of 2500 fT/cm in amplitude after noise reduction were rejected. Trials containing any remaining blinks were identified by individually visualizing raw activity for each epoch. If there was any sudden, stark increase of activity, the topography for that epoch was plotted. If the magnetic field pattern had the characteristic frontal distribution of a blink, that trial was also rejected. Additionally, trials corresponding to behavioral errors or response times within the length of our epochs were also excluded from further analyses. This procedure resulted in the exclusion of more than 85% of the trials for three participants due to excessive artifacts in their recordings (caused by construction work next door at the time of the experiment). Therefore, these three participants were excluded from further analyses. For the participants included in the analysis, the artifact and blink rejection routines resulted in the exclusion of 27.92% of the trials (11.9% SD), leaving 346.07 trials on average per subject (57.41 SD).

Data were averaged for each condition and subject. Averages were low-pass filtered at 40 Hz. Due to the excellent noise-conditions of the MEG facility, no high pass filtering was required. However, in the course of data analysis, it was observed that baseline correcting with the pre-stimulus interval created an artifactual sustained activity in the MEG averages starting at around 300 ms, due to the fact that the amplitude of the baseline period was higher than the evoked activity after early visual responses. The high baseline amplitude could have been due to task preparation, given that we had a blocked design where subjects needed to ready themselves to name the displays in the appropriate manner given the current task instruction. The drift introduced this way did not however qualitatively alter the obtained results: The same basic pattern was observed whether or not the data were baseline corrected with the pre-stimulus interval. Thus, to avoid visualizing the artifact, in the results reported below, the pre-stimulus interval was not used for baseline correction. Instead, for the purpose of defining the noise covariance matrix used in source analysis, we used an alternative approach offered by BESA Research 6.0, where baseline activity is calculated extracting the values of the 15% of all time points within our epochs that had the lowest global field power.[1]

To estimate the distributed electrical current image in the brain at each time sample we used the Minimum Norm Approach (Hämäläinen & Ilmoniemi, 1994) as implemented in BESA Research 6.0. The sources were evenly distributed using 1500 standard locations 10% and 30% below the smoothed standard brain surface (750 for each shell). The inverse solution problem was stabilized by the minimum norm mathematical constraint: Out of the many current distributions that could account for the recorded sensor data, the solution with the minimum L2 norm (i.e., the minimum total power of the current distribution) was used. This is performed as follows: First, the forward solution (leadfield matrix L) of all sources was calculated in the current head model. Then, the source activities S(t) of all source components were computed from the data matrix D(t) using an inverse regularized by the estimated noise covariance matrix $[S(t) = R \cdot L^T \cdot (L \cdot R \cdot L^T + C_N)^{-1} \cdot D(t)]$ where L is the leadfield matrix of the distributed regional source model, C_N denotes the noise correlation matrix in sensor space, and R is a weighting matrix in source space. The total activity of each regional source is computed as the root mean square of the source activities S(t) of its 3 (MEG:2) components. Additionally, we applied depth weighting and spatio-temporal weighting. Depth waiting was used in order for both deep and superficial sources to produce a similar, more focal result (as opposed to deep sources appearing very smeared in a minimum-norm reconstruction). This was computed by scaling the leadfield of each regional source with the largest singular value of the SVD (singular value decomposition) of the source's leadfield. The spatio-temporal weighting was conducted to assign large weight to the sources that are assumed to be more likely to contribute to the recorded data. We first divided the signal into a signal and a noise subspace. The correlation of the leadfield of a regional source i with the signal subspace (p_i) was computed to find out if the source location contributes to the measured data. The weighting matrix R then becomes a diagonal matrix. Each of the three (MEG: 2) components of a regional source get the same weighting value p_i^2 (Mosher and Leahy, 1998). There was no constraint posited on the dipole orientation (we used free orientation), the regularization constant was 1% and we did not apply any normalization (although we did use the residual variance fit criterion).[2] Regions of interest were defined in terms of Brodmann areas (BAs), which were isolated with the Talairach Daemon (Lancaster et al., 1997, 2000) from the BESA source space.

Statistical analysis

As the main goal of the current study was to compare combinatory effects in color modification as opposed to either complex number combinations or number quantification, we conducted a main analysis in the areas of the LATL (BAs 38, 20 and 21) that have previously been implicated in conceptual combination (Bemis and Pylkkänen, 2011; Pylkkänen et al., 2014; Westerlund and Pylkkänen, 2014). Although BA 20 and 21 stretch to more posterior regions of the temporal lobe, they were included in this analysis in order to cover anterior temporal cortex outside of the temporal pole (i.e., BA 38). In prior MEG studies, LATL combinatory effects have centered both around the pole (e.g., Del Prato and Pylkkänen, 2014; Pylkkänen et al., 2014) as well as more laterally (Westerlund and Pylkkänen, 2014; Westerlund et al., 2015). Crucially, we complemented our ROI-analyses with liberally thresholded whole brain contrasts capable of revealing the centers of activity within the ROIs.

In addition to the hypothesis-driven LATL analysis, we ran a separate more explorative analysis in areas that do not constitute the main focus of the study but could be sensitive to the current experimental manipulations. This second analysis included the angular gyrus (AG), the ventromedial prefrontal cortex (vmPFC) and the left inferior frontal gyrus (LIFG). The AG was included since we wanted to identify a possible locus for numerical quantification and plural composition, and this area has been previously reported to be involved both in number

[1] Please note that in both approaches the activity (noise or signal, respectively) is defined as root-mean-square across all respective latencies for each channel. The noise covariance matrix C_N is constructed as a diagonal matrix and the entries in the main diagonal are proportional to the noise activity of the individual channels (if selected) or are all equally proportional to the average noise activity over all channels. The noise covariance matrix C_N is then scaled such that the ratio of the Frobenius norms of the weighted leadfield projector matrix (L.R.L.T) and the noise covariance matrix C_N equals the Signal-to-Noise ratio.

[2] The source activities of the current solution were computed by multiplication of the inverse leadfield matrix times the measured data matrix. The leadfield matrix contains the topographies of the current dipoles, and is dependent on the current head model and results from the current dipole locations and orientations. After computing the source activities the leadfield matrix was multiplied by the source activity matrix to obtain the modeled data. The sum of the squared differences between the measured and the modeled data (summed over all channels and all samples) was divided by the power (the total variance) of the measured data. The result, called the residual variance, was minimized during the fit.

word processing (Dehaene et al., 2003) and plural representations (Boiteau et al., 2014; Carreiras et al., 2010; Domahs et al., 2012). The vmPFC was included as previous studies (Bemis and Pylkkänen, 2011, 2013; Pylkkänen et al., 2014) have found the vmPFC to be involved to some extent in basic composition. Following such studies, left and right BA11 were collapsed into a single ROI due to spatial adjacency along the midline. Lastly, although previous MEG studies on language production have not found composition effects in the LIFG (Del Prato and Pylkkänen, 2014; Pylkkänen et al., 2014) we also included this area in the analysis; given its general prominence in research in production (e.g., Haller et al., 2005; Indefrey et al., 2001a; Menenti et al., 2011). Due to the small number of regional sources within BA 44–45, they were also collapsed into a single ROI. BA 39 was used for the angular gyrus.

For the time-course data of each region, a non-parametric cluster permutation test (Maris and Oostenveld, 2007) with 10,000 permutations was used to identify temporal clusters during which the localized activity differed significantly between conditions, corrected for multiple comparisons over time. Prior to this test, MEG activity was averaged over all sources within a ROI. For initial cluster selection, we adopted the parameters of prior studies: 10 adjacent time points showing an effect at an alpha level of $p < 0.3$, (e.g., Bemis and Pylkkänen, 2011, 2012, 2013; Del Prato and Pylkkänen, 2014; Leiken and Pylkkänen, 2013; Pylkkänen et al., 2014; Westerlund and Pylkkänen, 2014). Then, for each cluster surviving these thresholds, a test statistic was constructed that was equal to the summed t-values of the point-by-point test-statistics over the selected cluster interval and finally, the cluster with the largest summed test statistic was chosen for further computations. Due to the last step, this test is only capable of identifying one effect within any given analysis interval and thus in order to be able to characterize potential earlier and later effects, all analyses were conducted both in an early (150–400 ms) and a late (400–600 ms) time window. Since all trials with faster reaction times than 700 ms were excluded from the analyses, we ensured these windows did not capture any late motion artifacts. For the largest cluster within an interval the corrected p-value ($p < .05$) was calculated as the ratio of permutations yielding a test statistic greater than the actual observed test statistic. Since only increases for combinatory conditions over lists were interpretable in light of our hypotheses and since our study in general was based on prior evidence that combinatory conditions elicit stronger activity than list conditions (Del Prato and Pylkkänen, 2014; Pylkkänen et al., 2014), all permutation t-tests were one-tailed. Finally, to moderately protect our analysis against false positives across multiple regions within the same analysis while still maintaining power, a false discovery rate (FDR; Benjamini and Hochberg, 1995; Genovese et al., 2002) of 0.1 was used throughout. This rate was kept somewhat liberal given that our main research question had no explorative component and the combinatory effects in our main dependent measure, LATL amplitude as measured by MEG, have already been replicated in numerous studies (Bemis and Pylkkänen, 2011, 2012, 2013; Westerlund and Pylkkänen, 2014; Del Prato and Pylkkänen, 2014; Pylkkänen et al., 2014; Leffel et al., 2014; Westerlund et al., 2015; Zhang and Pylkkänen, 2015).

Given that two of our LATL ROIs, BA 20 and 21, covered not only anterior but also posterior temporal cortex, as a final step in our analysis we visualized the centers of any obtained LATL effects with liberally thresholded uncorrected full brain analyses focused on the statistical peaks of the ROI effects in each comparison. Full brain contrasts were run at the time points of the highest uncorrected statistic in the ROI cluster and in the visualization, individual sources were plotted as red (indicating an increase for a combinatory condition) or blue (indicating a decrease for a combinatory condition) when they and at least two of their adjacent spatial and temporal neighbors showed an uncorrected significance of $p < .05$. As emphasized above, the purpose of these analyses was simply to address any potential spatial ambiguity in the ROI analyses.

Results

Behavioral results

Reaction times were submitted to a one way ANOVA with five levels. The results showed a main effect of condition [$F(4,96) = 84.75$, $p < .0001$]. Participants were the slowest naming the numbers in Color modification condition ($M = 940$ ms; $SD = 210$ ms) and planned t-tests showed that this delay was significant when compared to the other two experimental conditions: Numerical quantification [$t(24) = 3.22$, $p = .003$] and Complex Number naming [$t(24) = 10.5$, $p < .0001$]. Additionally, Numerical Quantification was also significantly slower than Complex Number naming [$t(24) = 12.83$, $p < .0001$]. There was no significant difference between our two control conditions; complex number list and number list [$t(24) = 0.83$, $p = .41$], although number list was slightly slower on average ($M = 788$ ms; $SD = 156$ ms for number list vs $M = 778$ ms $SD = 168$ ms for complex number list) (Fig. 2). Accuracy in all conditions was at ceiling, with each participant making an average of 6.3 errors in the course of the whole experiment (0.013%).

MEG results

In order to find the best control condition for our combinatorial effects, we ran a one way ANOVA with three levels (Complex number, Complex number list and Number list) in the LATL. Although non-significant, the result pattern showed a clear layered effect of composition, with Complex number eliciting the greatest activity followed by Complex number list, and Number list being the one eliciting the least activity (Additional Fig. 1). This pattern suggested that Complex number list condition may have elicited combinatorial activity to some extent. For this reason, we used Number list condition as the baseline control condition to assess combinatorial effects in our experimental manipulations.

The results of the pair wise comparisons conducted in the LATL (BA 38, 20, 21) revealed increased activation for Color modification over Number lists in all the analyzed areas (Fig. 3). The effect was reliable in BA20 [150–258 ms; $p = .04$] and in BA38 [184–349 ms; $p = .03$]. Additionally, a cluster of activity was also identified in BA21, although it did not reach significance [187–244 ms; $p = .18$]. Complex number condition also elicited significantly greater activity than Number lists in the three areas. Specifically, this increase in activity was reliable in BA21 [435–589 ms; $p = .02$], and marginally reliable in BA20 [505–590 ms; $p = .07$] and BA38 [513–583 ms; $p = .08$]. In contrast, the comparison between Numerical Quantification and Number List did not elicit significant differences in any of the analyzed areas (no clusters of activation were found in BA21 and BA38, and the cluster located in BA20 did not reach reliability [283–371 ms; $p = .1$]). To further contrast the differences between Complex number and Numerical Quantification, we ran an additional direct comparison between them and the results

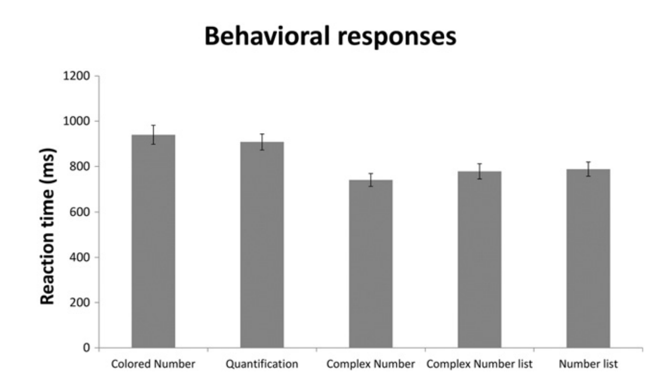

Fig. 2. Mean reaction times as a function of the performed naming task. Error bars show standard errors of the mean.

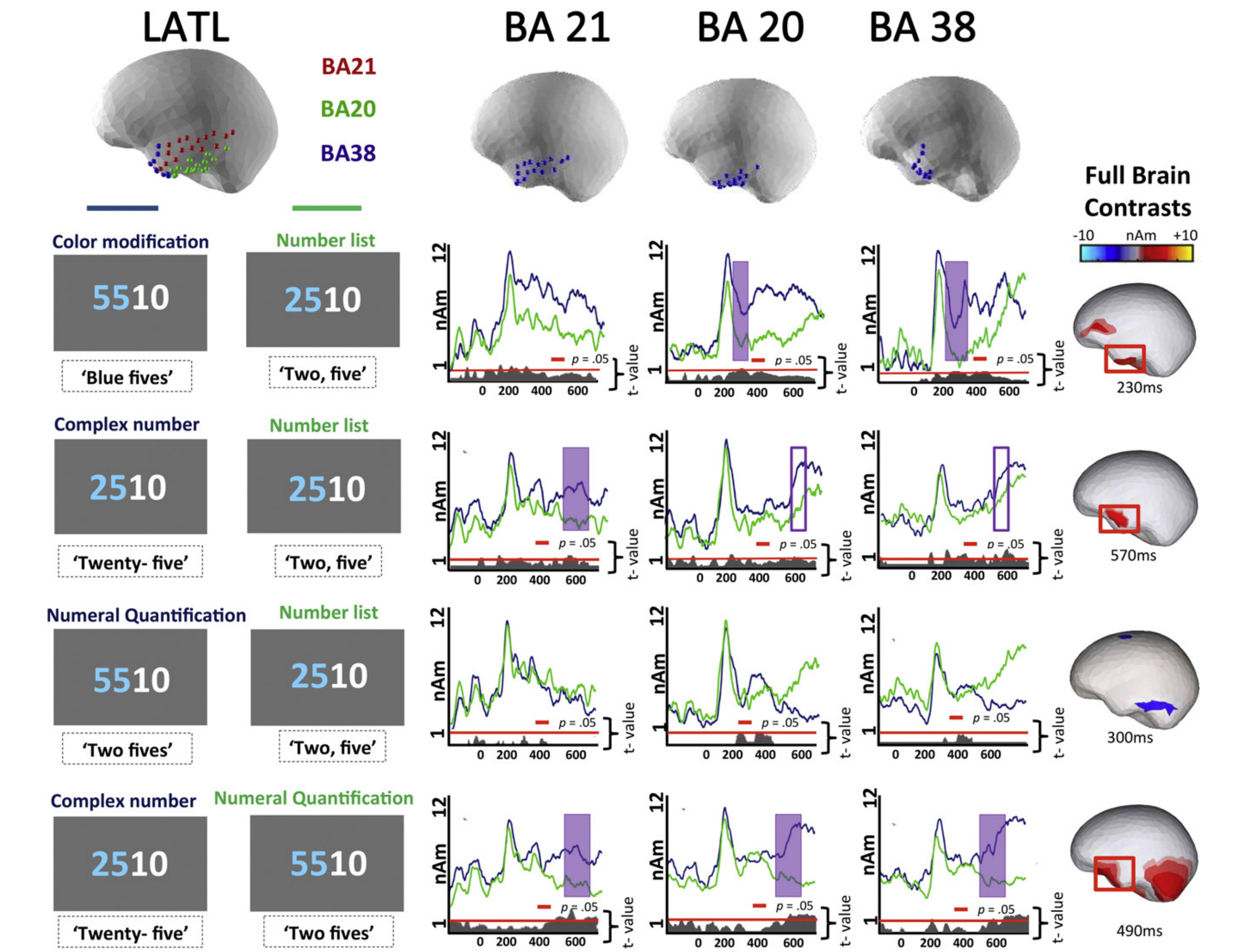

Fig. 3. ROI results for pair wise comparisons in the LATL, activation averaged across subjects. On the waveform plots, the shaded regions indicate that the difference in activity between the two tested conditions was significant at a $p = .05$ value (corrected), while the boxed region indicates marginally significant effects ($p < .1$). Significance was determined using a non-parametric, permutation test (Maris and Oostenveld, 2007) performed from 150 to 400 and 400 to 600 ms (10,000 permutations). The point-by-point t-statistic is also plotted in gray, with a red line indicating uncorrected significance at the $p < .05$ level. Finally, the right most panel visualizes the activity centers of the LATL effects obtained in the ROI analysis by plotting a source-by-source full brain comparison at the time of the statistical peak of the ROI effect. Individual sources are plotted as red when they and at least two of their adjacent spatial and temporal neighbors showed an increase for the combinatory condition at $p < .05$.

showed that Complex number elicited reliable increases in activity in BA38 [434–600 ms; $p = .003$], BA20 [437–600 ms; $p = .01$] and BA21 [400–600 ms; $p = .01$] (Fig. 3). Lastly, we also compared Color Modification and Complex number composition directly. These tests did not reveal any reliable differences between these conditions in either the earlier or the later time window.

Thus both complex number formation and color modification modulated the LATL, while numerical quantification did not affect it. When the center of this effect was visualized in the full brain contrasts, it localized ventrally in the anterior temporal cortex for color modification, while the center of the effect was a bit more superior for complex number composition (Fig. 3). Crucially though, both of these localizations are within the range of previously reported effect centers in prior LATL composition studies (Del Prato and Pylkkänen, 2014; Westerlund and Pylkkänen, 2014). Thus, there is not yet any detailed characterization as regards the precise locus of combinatory effects within the LATL, and given the somewhat fuzzy spatial resolution of MEG, it is also unclear whether such generalizations could be obtained with MEG.

The analyses on the AG aimed to test sensitivity for plurality in this area. Therefore, conditions containing plural NPs (Color modification and Numerical quantification) and singular NPs (Complex number and Number list) were contrasted. The results did not show significant differences for any of these contrasts. The vmPFC was reliably engaged when composing a complex number in comparison to naming a number list [458–600 ms; $p = .01$] but not when performing a numerical quantification (we did not find any cluster of activity) or performing a color modification [186–223 ms; $p = .22$]. Lastly, no comparisons elicited reliable clusters of activation in the LIFG (Additional Fig. 2).

Discussion

Complex number composition in the LATL

In this study we aimed to develop some understanding into the specifics of the combinatorial processes performed in the LATL and the nature of the elements that can enter these computations. Particularly, we tried to elucidate if the lack of LATL engagement during numerally quantified phrases such as *two cups* in previous production studies (Del Prato and Pylkkänen, 2014) was because number words are not a valid input element for the LATL or because the LATL is not involved in the computations underlying quantification in particular. To achieve this, we tested complex number expressions (*thirty-three*) where the combinatory mode intuitively resembles conceptual combination, already demonstrated to drive the LATL (comprehension: Baron and Osherson, 2011; Bemis and Pylkkänen, 2011, 2013; Westerlund and Pylkkänen, 2014; production: Del Prato and Pylkkänen, 2014; Pylkkänen et al., 2014), and compared this to numerical quantification. According to our results, the reason for the lack of LATL effects in numerical quantifications is not the numerical input but rather the mode of composition: when numbers compose into complex concepts as in *thirty-two*, LATL activity is increased in comparison to both non-combinatory lists (*three*, *two*) as well as to quantificational phrases (*three twos*). Thus our findings suggest that the LATL is not a general purpose combiner of meanings but rather specializes in some version of conceptual combination, potentially delimited to situations where one combining element characterizes a property of the other.

Additionally, the finding of combinatorial activity for our complex number condition conforms to theories where complex numbers undergo a composition process before being produced (Deloche and Seron, 1987), as opposed to being processed holistically (Brysbaert, 1995; Dehaene et al., 1990; Reynvoet and Brysbaert, 1999). In other words, our results suggest that in order to produce "twenty-three", the concept of "twenty" and the concept of "three" are retrieved and combined. This proposal is consistent with McCloskey et al.'s (1986) model where complex number production involves the generation of a syntactic frame that specifies each to-be-retrieved word in terms of a number-lexical class (ones, teens…) and a position within that class. Moreover, it conforms to proposals within theoretical linguistics where complex number composition follows the standard principles of semantic and syntactic combination (Ionin and Matushansky, 2006).

Despite the anatomical overlap in regions activated by complex number naming and color modification, we did observe an unpredicted difference in the time course of these effects: Composition effects for complex numbers occurred around 200 ms later than the adjectival modification effects. This difference is especially interesting given that articulation onsets exhibited the reverse pattern: complex number productions were faster than color modifications (though the plural morphology of the color modification likely contributed to this; see next section). This makes explanations in terms of any type of general difficulty somewhat unlikely: for example, the two conditions differed in the number of syllables, but under a syllable-based account, it would be unclear how faster reaction times could be elicited when more syllables are articulated. Instead, inspection of the source waveforms (Fig. 3) suggests that in general, similar peaks and valleys are elicited both for color modification and complex number formation, but along this waveform morphology, the two conditions show significant effects at different stages. Given that language production is thought to involve a complex interplay of both comprehension and production processes (e.g., Pickering and Garrod, 2013), one could then speculate that the combinatory processing of complex numbers occurs most strongly at a later stage in this sequence than the combinatory processing of adjectival modifications. Thus in future work it will be possible to articulate and test more specific hypotheses regarding such a contrast.

Contributions of the AG, vmPFC and LIFG

The AG was included in our analysis to test whether we would replicate the finding by Domahs et al. (2012) that plural nouns elicit reliably greater AG activity than singular nouns. We tested this by comparing the two conditions that involved naming a plural NP (Color modification and Numerical Quantification) to the two that involved a singular NP (Complex number and Number list), but did not find any reliable differences between conditions. Interestingly, Domahs et al.'s study did not find the AG to be uniquely involved in the processing of plural entities per se, but rather in the processing of non-singular items (including mass nouns). Thus, a possible reason for the lack of differences between our conditions could be that all of our experimental conditions, including the singular NPs, were composed of number words which represented non-singular concepts. This interpretation of the AG as a non-singularity representation hub but not a syntactic number marking region is also convergent with Carreiras et al. (2010), who did not find the AG to be involved in number agreement processing.

Interestingly, in the behavioral data, we observed longer naming latencies in exactly the two conditions that involved naming a plural NP (Color modification and Numerical Quantification) as compared to the three conditions that involved naming singular NPs. This "plural effect" conforms to prior findings in comprehension showing delayed responses to plural as compared to singular NPs (Tucker et al., 2015; Lau et al., 2008; Wagers et al., 2009), suggesting that the processing of syntactically plural entities is more effortful than singular ones. However, our results do not speak to the neural substrates of this phenomenon; thus, future work will need to unveil the correlates of syntactic number marking and will have to establish the relationship between grammatical number case and general number magnitude processing.

The analyses of vmPFC activity revealed trends towards increases for complex numbers and color modifications over number list whereas quantification did not elicit higher activity than the list. These results provide moderate support for an implication of the vmPFC in basic combination (Bemis and Pylkkänen, 2011; Pylkkänen et al., 2014) and suggest that numerical input can also drive combinatorial activity in the vmPFC.

Finally, consistent with prior MEG studies on basic combinatory phrases both in comprehension (Bemis and Pylkkänen, 2011, 2012) and production (Del Prato and Pylkkänen, 2014; Pylkkänen et al., 2014), we did not observe combinatory effects in the LIFG. Thus to the extent that the LIFG may be involved in linguistic composition (e.g., Hagoort, 2005), the nature of its contribution is not a time-locked evoked response of the sort that would be revealed by the type of source modeling of averaged data as performed here.

Conclusion

In this work we set out to delineate the combinatory computation housed in the LATL, with a focus on number words and their composition. We found that the LATL composes number words but only if the composing items combine into a more complex concept (*thirty two*) and not when one word enumerates the number of tokens of the other (*three twos*). Our findings suggest that the engagement of the LATL is determined by the computations underlying the performed combinatorial process as opposed to the nature of the input items.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.neuroimage.2015.08.049.

Acknowledgments

This research was supported by the National Science Foundation grant BCS-1221723 (LP) and grant G1001 from the NYUAD Institute, New York University Abu Dhabi (LP).

References

Adachi, Y., Shimogawara, M., Higuchi, M., Haruta, Y., Ochiai, M., 2001. Reduction of non-periodic environmental magnetic noise in MEG measurement by continuously adjusted least squares method. IEEE Trans. Appl. Supercond. 11, 669–672.
Alario, F.X., Costa, A., Caramazza, A., 2002. Frequency effects in noun phrase production: implications for models of lexical access. Lang. Cogn. Process. 17 (3), 299–319.
Bale, A.C., Khanjian, H., 2011. Cross-Linguistic Representations of Numerals and Number Marking. Proceedings of SALT vol. 20, pp. 582–598.
Balota, D.A., Yap, M.J., Hutchison, K.A., Cortese, M.J., Kessler, B., Loftis, B., Neely, J.H., Nelson, D.L., Simpson, G.B., Treiman, R., 2007. The English Lexicon Project. Behav. Res. Methods 39 (3), 445–459.
Baron, S.G., Osherson, D., 2011. Evidence for conceptual combination in the left anterior temporal lobe. NeuroImage 55 (4), 1847–1852.
Barwise, J., Cooper, R., 1981. Generalized quantifiers and natural language. Linguist. philos. 4 (2), 159–219.
Bemis, D.K., Pylkkänen, L., 2011. Simple composition: a magnetoencephalography investigation into the comprehension of minimal linguistic phrases. J. Neurosci. 31 (8), 2801.

Bemis, D.K., Pylkkänen, L., 2012. Basic linguistic composition recruits the left anterior temporal lobe and left angular gyrus during both listening and reading. Cereb. Cortex 23 (8), 1859–1873.
Bemis, D.K., Pylkkänen, L., 2013. Basic linguistic composition recruits the left anterior temporal lobe and left angular gyrus during both listening and reading. Cereb. Cortex 23 (8), 1859–1873.
Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. Ser. B Methodol. 289–300.
Bennett, M.R., 1975. Some Extensions of a Montague Fragment of English. vol. 129. Indiana University Linguistics Club.
Binder, J.R., Desai, R.H., 2011. The neurobiology of semantic memory. Trends Cogn. Sci. 15 (11), 527–536.
Binder, J.R., Desai, R.H., Graves, W.W., Conant, L.L., 2009. Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. Cereb. Cortex 19 (12), 2767–2796.
Boiteau, T.W., Bowers, E., Nair, V.A., Almor, A., 2014. The neural representation of plural discourse entities. Brain Lang. 137, 130–141.
Bonner, M.F., Peelle, J.E., Cook, P.A., Grossman, M., 2013. Heteromodal conceptual processing in the angular gyrus. NeuroImage 71, 175–186.
Brennan, J., Pylkkänen, L., 2012. The time-course and spatial distribution of brain activity associated with sentence processing. NeuroImage 60 (2), 1139–1148.
Bright, P., Moss, H., Tyler, L., 2004. Unitary vs multiple semantics: PET studies of word and picture processing. Brain Lang. 89 (3), 417–432.
Brysbaert, M., 1995. Arabic number reading: on the nature of the numerical scale and the origin of phonological recoding. J. Exp. Psychol. Gen. 124, 434–452.
Carreiras, M., Carr, L., Barber, H.A., Hernandez, A., 2010. Where syntax meets math: right intraparietal sulcus activation in response to grammatical number agreement violations. NeuroImage 49 (2), 1741–1749.
Clarke, A., Taylor, K.I., Tyler, L.K., 2011. The evolution of meaning: spatio-temporal dynamics of visual object recognition. J. Cogn. Neurosci. 23 (8), 1887–1899.
Clarke, A., Taylor, K.I., Devereux, B., Randall, B., Tyler, L.K., 2013. From perception to conception: how meaningful objects are processed over time. Cereb. Cortex 23 (1), 187–197.
Dehaene, S., 1992. Varieties of numerical abilities. Cognition 44 (1), 1–42.
Dehaene, S., Dupoux, E., Mehler, J., 1990. Is numerical comparison digital? Analogical and symbolic effects in two-digit number comparison. J. Exp. Psychol. Hum. Percept. Perform. 16, 626–641.
Dehaene, S., Piazza, M., Pinel, P., Cohen, L., 2003. Three parietal circuits for number processing. Cogn. Neuropsychol. 20 (3–6), 487–506.
Del Prato, P., Pylkkänen, L., 2014. MEG evidence for conceptual combination but not numeral quantification in the left anterior temporal lobe during language production. Front. Psychol. 5 (524).
Deloche, G., Seron, X., 1987. Numerical Transcoding: A General Production Model. Lawrence Erlbaum Associates, Inc.
Domahs, F., Nagels, A., Domahs, U., Whitney, C., Wiese, R., Kircher, T., 2012. Where the mass counts: common cortical activation for different kinds of nonsingularity. J. Cogn. Neurosci. 24 (4), 915–932.
Dotan, D., Friedmann, N., Dehaene, S., 2014. Breaking down number syntax: spared comprehension of multi-digit numbers in a patient with impaired digit-to-word conversion. Cortex 59, 62–73.
Friederici, A.D., 2011. The brain basis of language processing: from structure to function. Physiol. Rev. 91 (4), 1357–1392.
Friederici, A.D., Meyer, M., Von Cramon, D.Y., 2000. Auditory language comprehension: an event-related fMRI study on the processing of syntactic and lexical information. Brain Lang. 74 (2), 289–300.
Gauthier, I., Anderson, A.W., Tarr, M.J., Skudlarski, P., Gore, J.C., 1997. Levels of categorization in visual recognition studied using functional magnetic resonance imaging. Curr. Biol. 7 (9), 645–651.
Genovese, C.R., Lazar, N.A., Nichols, T., 2002. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. NeuroImage 15 (2), 870–878.
Grabowski, T.J., Damasio, H., Tranel, D., Ponto, L.L.B., Hichwa, R.D., Damasio, A.R., 2001. A role for left temporal pole in the retrieval of words for unique entities. Hum. Brain Mapp. 13 (4), 199–212.
Grossberg, S., Repin, D.V., 2003. A neural model of how the brain represents and compares multi-digit numbers: spatial and categorical processes. Neural Netw. 16 (8), 1107–1140.
Hagoort, P., 2005. On Broca, brain, and binding: a new framework. Trends Cogn. Sci. 9 (9), 416–423.
Hagoort, P., Indefrey, P., 2014. The neurobiology of language beyond single words. Annu. Rev. Neurosci. 37, 347–362.
Haller, S., Radue, E.W., Erb, M., Grodd, W., Kircher, T., 2005. Overt sentence production in event-related fMRI. Neuropsychologia 43, 807–814.
Hämäläinen, M.S., Ilmoniemi, R.J., 1994. Interpreting magnetic fields of the brain: minimum norm estimates. Med. Biol. Eng. Comput. 32 (1), 35–42.
Hampton, J.A., 1997. Conceptual combination. In: Lamberts, K., Shanks, D. (Eds.), Knowledge, Concepts, and Categories. UCL Press, London, pp. 133–159.
Heim, S., Eickhoff, S.B., Amunts, K., 2008. Specialisation in Broca's region for semantic, phonological, and syntactic fluency? NeuroImage 40 (3), 1362–1368.
Humphries, C., Binder, J.R., Medler, D.A., Liebenthal, E., 2006. Syntactic and semantic modulation of neural activity during auditory sentence comprehension. J. Cogn. Neurosci. 18 (4), 665–679.
Humphries, C., Binder, J.R., Medler, D.A., Liebenthal, E., 2007. Time course of semantic processes during sentence comprehension: an fMRI study. NeuroImage 36 (3), 924–932.
Hurford, J., 1975. The Linguistic Theory of Numerals. Cambridge University Press, Cambridge.
Hurford, J., 1987. Language and Number: The Emergence of a Cognitive System. Blackwell, Oxford.
Hurford, J., 2001. Numeral systems. In: Smelser, N.J., Baltes, P.B. (Eds.), International Encyclopedia of the Social and Behavioral Sciences. Pergamon, Amsterdam, pp. 10756–10761.
Hurford, J., 2003. The Interaction Between Numerals and Nouns. In: Plank, F. (Ed.), Noun Phrase Structure in the Languages of Europe. Typology of Languages in Europe 7. Mouton de Gruyter. The Hague.
Indefrey, P., 2012. Hemodynamic studies of syntactic processing. In: Faust, M. (Ed.), The Handbook of the Neuropsychology of Language. MA: Blackwell, Malden, pp. 209–228.
Indefrey, P., Brown, C.M., Hellwig, F., Amunts, K., Herzog, H., Hagoort, P., 2001a. A neural correlate of syntactic encoding during speech production. Proc. Natl. Acad. Sci. 98, 5933–5936.
Indefrey, P., Hagoort, P., Herzog, H., Seitz, R.J., Brown, C.M., 2001b. Syntactic processing in left prefrontal cortex is independent of lexical meaning. NeuroImage 14 (3), 546–555.
Ionin, T., Matushansky, O., 2006. The composition of complex cardinals. J. Semant. 23 (4), 315–360.
Kaufman, E.L., Lord, M.W., Reese, T.W., Volkmann, J., 1949. The discrimination of visual number. Am. J. Psychol. 62, 498–525.
Lancaster, J.L., Rainey, L.H., Summerlin, J.L., Freitas, C.S., Fox, P.T., Evans, A.C., ... Mazziotta, J.C., 1997. Automated labeling of the human brain: a preliminary report on the development and evaluation of a forward-transform method. Hum. Brain Mapp. 5 (4), 238.
Lancaster, J.L., Woldorff, M.G., Parsons, L.M., Liotti, M., Freitas, C.S., Rainey, L., ... Fox, P.T., 2000. Automated Talairach atlas labels for functional brain mapping. Hum. Brain Mapp. 10 (3), 120–131.
Lau, E.F., Phillips, C., Poeppel, D., 2008. A cortical network for semantics:(de) constructing the N400. Nat. Rev. Neurosci. 9 (12), 920–933.
Leffel, T., Lauter, M., Westerlund, M., Pylkkänen, L., 2014. Restrictive vs. non-restrictive composition: a magnetoencephalography study. Lang. Cogn. Neurosci. 29 (10), 1191–1204.
Leiken, K., Pylkkänen, L., 2013. MEG evidence that the LIFG effect of object extraction requires similarity-based interference. Language and Cognitive Processes.
Link, G., 2002. The logical analysis of plurals and mass terms: a lattice-theoretical approach. Formal Semantics: The Essential Readingspp. 127–146.
Maris, E., Oostenveld, R., 2007. Nonparametric statistical testing of EEG- and MEG-data. J. Neurosci. Methods 164 (1), 177–190.
Mazoyer, B.M., Tzourio, N., Frak, V., Syrota, A., Murayama, N., Levrier, O., et al., 1993. The cortical representation of speech. J. Cogn. Neurosci. 5 (4), 467–479.
McCloskey, M., 1992. Cognitive mechanisms in numerical processing: evidence from acquired dyscalculia. Cognition 44 (1), 107–157.
McCloskey, M., Caramazza, A., Basili, A., 1985. Cognitive mechanisms in number processing and calculation: evidence from dyscalculia. Brain Cogn. 4 (2), 171–196.
McCloskey, M., Sokol, S.M., Goodman, R.A., 1986. Cognitive processes in verbal-number production: Inferences from the performance of brain-damaged subjects. J. Exp. Psychol. Gen. 115 (4), 307.
Medin, D.L., Shoben, E.J., 1988. Context and structure in conceptual combination. Cogn. Psychol. 20 (2), 158–190.
Menenti, L., Gierhan, S.M.E., Segaert, K., Hagoort, P., 2011. Shared language: overlap and segregation of the neuronal infrastructure for speaking and listening revealed by functional MRI. Psychol. Sci. 22, 1173–1182.
Mesulam, M.M., 2000. Behavioral neuroanatomy: large-scale networks, association cortex, frontal syndromes, the limbic system, and hemispheric specializations. Princ. Behav. Cogn. Neurol. 2, 1–120.
Meyer, A.S., 1996. Lexical access in phrase and sentence production: results from picture–word interference experiments. J. Mem. Lang. 35, 477–496.
Montague, R., 1974. Formal Philosophy; Selected Papers of Richard Montague.
Mosher, J.C., Leahy, R.M., 1998. Recursive MUSIC: a framework for EEG and MEG source localization. Biomed. Eng.IEEE Trans. 45 (11), 1342–1354.
Murphy, G.L., 1990. Noun phrase interpretation and conceptual combination. J. Mem. Lang. 29 (3), 259–288.
Nuerk, H.C., Weger, U., Willmes, K., 2001. Decade breaks in the mental number line? Putting tens and units back into different bins. Cognition 82, B25–B33.
Pallier, C., Devauchelle, A.D., Dehaene, S., 2011. Cortical representation of the constituent structure of sentences. Proc. Natl. Acad. Sci. 108 (6), 2522.
Pandya, D.N., Seltzer, B., 1982. Association areas of the cerebral cortex. Trends Neurosci. 5, 386–390.
Pickering, M.J., Garrod, S., 2013. An integrated theory of language production and comprehension. Behav. Brain Sci. 36 (04), 329–347.
Price, A.R., Bonner, M.F., Peelle, J.E., Grossman, M., 2015. Converging evidence for the neuroanatomic basis of combinatorial semantics in the angular gyrus. J. Neurosci. 35 (7), 3276–3284.
Pylkkänen, L., Bemis, D.K., Blanco Elorrieta, E., 2014. Building phrases in language production: an MEG study of simple composition. Cognition 133, 371–384.
Reynvoet, B., Brysbaert, M., 1999. Single-digit and two-digit Arabic numbers address the same semantic number line. Cognition 72, 191–201.
Rogalsky, C., Hickok, G., 2009. Selective attention to semantic and syntactic features modulates sentence processing networks in anterior temporal cortex. Cereb. Cortex 19 (4), 786.
Rogers, T.T., Hocking, J., Noppeney, U., Mechelli, A., Gorno-Tempini, M.L., Patterson, K., et al., 2006. Anterior temporal cortex and semantic memory: reconciling findings from neuropsychology and functional imaging. Cogn. Affect. Behav. Neurosci. 6 (3), 201–213.
Saltzman, I.J., Gamer, W.R., 1948. Reaction time as a measure of span of attention. Am. J. Psychol. 25, 227–241.
Scha, R., 1984. Distributive, Collective and Cumulative Quantification. Truth, Interpretation, and Information, GRASS 2 pp. 131–158.

Schriefers, H., De Ruiter, J.P., Steigerwald, M., 1999. Parallelism in the production of noun phrases: experiments and reaction time models. J. Exp. Psychol. Learn. Mem. Cogn. 25 (3), 702.

Segaert, K., Menenti, L., Weber, K., Petersson, K.M., Hagoort, P., 2012. Shared syntax in language production and language comprehension—an fMRI study. Cereb. Cortex 22 (7), 1662–1670.

Stowe, L.A., Broere, C.A.J., Paans, A.M.J., Wijers, A.A., Mulder, G., Vaalburg, W., et al., 1998. Localizing components of a complex task: sentence processing and working memory. NeuroReport 9 (13), 2995.

Stromswold, K., Caplan, D., Alpert, N., Rauch, S., 1996. Localization of syntactic comprehension by positron emission tomography. Brain Lang. 52 (3), 452–473.

Thompson-Schill, S.L., D'Esposito, M., Aguirre, G.K., Farah, M.J., 1997. Role of left inferior prefrontal cortex in retrieval of semantic knowledge: a reevaluation. Proc. Natl. Acad. Sci. U. S. A. 94, 14792–14797.

Tucker, M.A., Idrissi, A., Almeida, D., 2015. Representing number in the real-time processing of agreement: self-paced reading evidence from Arabic. Front. Psychol. 6, 314.

Tyler, L.K., Stamatakis, E.A., Bright, P., Acres, K., Abdallah, S., Rodd, J., et al., 2004. Processing objects at different levels of specificity. J. Cogn. Neurosci. 16 (3), 351–362.

Tyler, L.K., Marslen-Wilson, W.D., Randall, B., Wright, P., Devereux, B.J., Zhuang, J., ... Stamatakis, E.A., 2011. Left inferior frontal cortex and syntax: function, structure and behaviour in patients with left hemisphere damage. Brain 134 (2), 415–431.

Vandenberghe, R., Nobre, A.C., Price, C.J., 2002. The response of left temporal cortex to sentences. J. Cogn. Neurosci. 14 (4), 550–560.

Verguts, T., De Moor, W., 2005. Two-digit comparison: decomposed, holistic, or hybrid? Exp. Psychol. 52 (3), 195–200.

Wagers, M.W., Lau, E.F., Phillips, C., 2009. Agreement attraction in comprehension: representations and processes. J. Mem. Lang. 61 (2), 206–237.

Westerlund, M., Pylkkänen, L., 2014. The role of the left anterior temporal lobe in semantic composition vs. semantic memory. Neuropsychologia 57, 59–70.

Westerlund, M., Kastner, I., Al Kaabi, M., Pylkkänen, L., 2015. The LATL as locus of composition: MEG evidence from English and Arabic. Brain Lang. 141, 124–134.

Wisniewski, E.J., 1996. Construal and similarity in conceptual combination. J. Mem. Lang. 35 (3), 434–453.

Xu, J., Kemeny, S., Park, G., Frattali, C., Braun, A., 2005. Language in context: emergent features of word, sentence, and narrative comprehension. NeuroImage 25 (3), 1002–1015.

Yeterian, E.H., Pandya, D.N., 1985. Corticothalamic connections of the posterior parietal cortex in the rhesus monkey. J. Comp. Neurol. 237 (3), 408–426.

Zhang, L., Pylkkänen, L., 2015. The Interplay of Composition and Concept Specificity in the Left Anterior Temporal Lobe: An MEG Study. NeuroImage 111, 228–240.

Zhou, X., Chen, C., Chen, L., Dong, Q., 2008. Holistic or compositional representation of two-digit numbers? Evidence from the distance, magnitude, and SNARC effects in a number-matching task. Cognition 106 (3), 1525–1536.